Large Language Models as Commonsense Knowledge for Large-Scale Task Planning

Year: 2023 | Citations: 311 | Authors: Zirui Zhao, W. Lee, David Hsu

---

Abstract

Large-scale task planning is a major challenge. Recent work exploits large language models (LLMs) directly as a policy and shows surprisingly interesting results. This paper shows that LLMs provide a commonsense model of the world in addition to a policy that acts on it. The world model and the policy can be combined in a search algorithm, such as Monte Carlo Tree Search (MCTS), to scale up task planning. In our new LLM-MCTS algorithm, the LLM-induced world model provides a commonsense prior belief for MCTS to achieve effective reasoning; the LLM-induced policy acts as a heuristic to guide the search, vastly improving search efficiency. Experiments show that LLM-MCTS outperforms both MCTS alone and policies induced by LLMs (GPT2 and GPT3.5) by a wide margin, for complex, novel tasks. Further experiments and analyses on multiple tasks -- multiplication, multi-hop travel planning, object rearrangement -- suggest minimum description length (MDL) as a general guiding principle: if the description length of the world model is substantially smaller than that of the policy, using LLM as a world model for model-based planning is likely better than using LLM solely as a policy.